
Mathematical Methods for Computer Science I

Fall 2016

Outline 12

2.8. Context-free grammars and languages.

Definition: A **context-free grammar** (CFG) is a tuple $G = (V, T, P, S)$ where V and T are finite sets of **variables** and **terminals**. Terminals are symbols (letters) which cannot be modified. P are **productions** or **rules** and S is the **start symbol**.

Definition: A language L generated by a context-free grammar is a **context-free language** (CFL).

Definition: A context-free grammar is in **Chomsky normal form** if all productions either have the form $A \rightarrow a$ or $A \rightarrow BC$ with $a \in T$ and $A, B, C \in V$.

Definition: A **pushdown automaton** (PDA) is a 6-tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$ (Γ is the set of stack symbols) with $\delta : Q \times (\Sigma \cup \{\varepsilon\}) \times (\Gamma \cup \{\varepsilon\}) \rightarrow Q \times (\Gamma \cup \{\varepsilon\})$

A pushdown automaton thus is an ε -NFA with a stack to store symbols, and

-) last in, first out access of information
-) unlimited memory: the stack can save an infinite amount of information
-) writing a symbol on the stack is called "pushing" that symbol
-) removing a symbol from the stack is called "popping" the symbol

Theorem. *A language is context-free \iff a PDA recognizes it.*

Corollary. *Every regular language is also context-free.*

Theorem. *If L is a regular language, then there is a deterministic PDA which recognizes it.*

Lemma (Pumping lemma for context-free languages). *Let L be a CFL. Then there is a constant n such that if $z \in L$ with $|z| \geq n$, then $z = uvwxy$ s.t. :*

1. $|vwx| \leq n$
2. $vx \neq \varepsilon$ (that is v and x are not simultaneously ε)
3. $uv^iwx^iy \in L$ for all $i \geq 0$. That is v and x are pumped simultaneously.

2.9. Context-sensitive grammars and languages.

Definition: A **context-sensitive grammar** (CSG) $G = (V, T, P, S)$ has the same definition as a context-free grammar, except that the left-hand side ℓ of a production $\ell \rightarrow r$ can contain any combination of variables (nonterminals) and terminals with the only restriction that $|\ell| \leq |r|$.

Definition: A language L generated by a context-sensitive grammar is a **context-sensitive language** (CSL).